Jialuo Liu

http://liujl93.github.io/

Email : liujl@rams.colostate.edu
Mobile : +1-970-889-6728

Summary

- Award-winning Ph.D. researcher with solid theoretical and applied background. Experienced in solving complex quantitative problems and developing efficient algorithms in spatial/spatio-temporal statistics.
- Hands-on experience in dealing with large datasets, performing data cleaning/visualization, feature engineering and applying machine learning algorithms to solve business problems in `R/Python`.
- Enthusiastic team-player, gained from 4 years of collaboration with researchers from different professional background.

Education

- **Colorado State University** Fort Collins, CO
  Doctor of Philosophy (Ph.D.). in Statistics; GPA: 4.00/4.00 *Aug. 2014 – May. 2019 (Expected)*
  - **Dissertation:** 'Statistical Modeling and Inference for Spatial and Spatio-temporal Data'
  - **Awards**: **James L., M. Leslie, & Edna Madison Memorial Award**, elected by statistics faculty as the **outstanding** graduate student **(top 1)** in the department. **Franklin A. Graybill Linear Models Award**, awarded annually to a **top** graduate student in **linear models**.

- **University of Science and Technology of China** Hefei, China
  Bachelor of Science (B.S.). in Statistics. *Sept. 2010 – June. 2014*

Projects

- **Semiparametric Modeling and Bandwidth Selection** *R, Kernel Regression, Cross Validation* *2015-2018*
  - Developed a new spatio-temporal methodology that integrates multiple data sources for data processing.
  - Proposed a novel bandwidth selection procedure based on bimodal kernel functions for kernel regression to address the breakdown of cross-validation in the analysis of correlated data.
  - Conducted massive simulations on distributed systems in parallel to explore the performances of the method.
  - Created an R package `STplm` to automate aforementioned modeling and bandwidth selection procedure.

- **Krylov Subspace Method for Large Spatial Datasets** *R, Conjugate Gradient, Gaussian Process* *2017-2019*
  - Proposed an approximation to the Gaussian log-likelihood function using Krylov subspace methods.
  - Implemented the conjugate gradient method to solve large linear systems.
  - Presented a stochastic estimator based on Monte Carlo method and Gauss quadrature rule to approximate the log-determinant.

- **House Price Prediction by Stacked Regressions** *Python, Regression, Stacking* *2019*
  - Applied supervised machine learning to house price prediction.
  - Performed data visualization, missing value imputation, variable selection, cross validation, and parameter tuning.
  - Built a Stacked model on top of XGBoost, ElasticNet, Kernel Ridge regression and Lasso models. Achieved 11.56% Root Mean Squared Error, ranked top 13% in the competition.

- **Species Diversity and Abundance of Spiders** *R, Data Cleaning/Visualization, Pattern Analysis* *2017-2018*
  - Worked closely with ecologists to clean, summarize and visualize ecological data using dplyr and ggplot2 in R.
  - Assessed the statistical significance of the relationship of species occurrence/abundance and groups of sites by multi-level pattern analysis.

Skills

- **Languages**: R, Python, SAS, MySQL, C/C++.
- **Computing**: Parallel computing on high performance computing systems; Advanced programming in R (`Rcpp` for R and C++ Integration and `Rmpi` for message passing.) Hands-on Big Data experience with **Spark** and **MLlib**.
- **Miscellaneous**: LaTeX, Git, Markdown, shell scripting, Linux.
- **Relevant Coursework**: probability theory, mathematical statistics, data analysis and regression, experimental design, mixed models, time series, machine learning, optimization, and nonlinear programming.

Experience

- **Colorado State University** Fort Collins, CO

  Statistical Consultant *2015 - 2016*

  - Analyzed complex data sets from clients to create clear and compelling reports and visualizations. Cleaned data, provided software support (e.g., R, SAS, and JMP), and evaluated the results for the clients.

  Online Course Coordinator *2015 - 2018*

  - Provided course specific advice to master students on a variety of graduate-level courses including experimental design, mixed models, statistical learning and data mining.

  Teaching Assistant *2017 - 2018*

  - Independently instructed the course 'General Statistics', delivered a range of teaching and assessment activities including tutorials directed towards the delivery of subjects at undergraduate level.

Publications

- **Liu, J.**, Chu, T., Zhu, J., & Wang, H. "On Locally Stationary Spatio-temporal Processes and Inference for Irregularly Spaced and Timed Data". In progress.
- **Liu, J.**, Chu, T., Zhu, J., & Wang, H. "Krylov Subspace Method for Large Spatial Datasets". In progress.
- Knutson, E. M, **Liu, J.**, and B. C. Kondratieff. "Insect Occupants of Dominant Shrub Communities of Pawnee and Comanche National Grasslands, Colorado". Accepted.